

Code Breaking for Automatic Speech Recognition

Frederick Jelinek

Center for Language and Speech Processing,
The Johns Hopkins University, Baltimore, USA

Abstract. Practical automatic speech recognition is of necessity a (near) real time activity performed by a system whose structure is fixed and whose parameters once trained may be adapted on the basis of the speech that the system observed during recognition.

However, in specially important situations (e.g., recovery of out-of-vocabulary words) the recognition task could be viewed as an activity akin to code-breaking to whose accomplishment can be devoted an essentially infinite amount effort. In such a case everything would be fair, including, for instance, the retraining of a language and/or acoustic model on the basis of newly acquired data (from the Internet!) or even a complete change of the recognizer paradigm.

An obvious way to proceed is to use the basic recognizer to produce a lattice or confusion network and then do the utmost to eliminate ambiguity. Another possibility is to create a list of frequent confusions (for instance the pair IN and AND) and prepare an appropriate individual decision processes to resolve each when it occurs in test data. We will report on our initial code breaking effort.